ES3N: A Semantic Approach to Data Management in Sensor Networks

Micah Lewis, Delroy Cameron, Shaohua Xie, I. Budak Arpinar

Computer Science Department
University of Georgia
Athens, GA 30602

Outline

- Background
- Problem
- ES3N Implementation
- Semantic Benefits
- Future Work
- Demo
Background

- Cargill Industries Inc.
  - Corporate Headquarters in Minneapolis, MN
  - Origin as family-owned food business
  - Started in 1865
  - 149,000 employees
  - Spans 63 countries

Cargill

- International provider of food services
- Commercial cereal grain and oil seed storage
- Food Processing (soybean and corn)
- Provide quality product to consumers
Background…

- USDA ARS NPRL
  - USDA created 1862
  - ARS created 1953
  - NPRL established 1965

NPRL

- Subsidiary research unit of the ARS
- Unshelled and shelled peanut research
- Quality for pre and post harvest
- Control aflatoxin
Problem

- Cargill
  - Primitive data acquisition
  - No data storage mechanism
  - No possibility for data analysis or mining

Problem...

- NPRL
  - Data management
  - Laborious human analysis
  - Query functionality nonexistent
ES3N Implementation

- Three targeted areas:
  - Data acquisition
  - Data Storage
  - Data Management

---

ES3N Implementation

- Four main components:
  - Sensor Network
  - Data Analysis and Query Processing Unit
  - Ontology
  - GUI (Graphical User Interface)
Development

- Data collection
- Memory caching
- Data Tagging
- Ontology representation
- Query processing
Data Collection

- Raw data retrieved from sensors
- Ontology provides persistent storage
- Data reside in ontology in RDF files

Memory Caching

- Preserve efficiency of system
- Effectively manage memory
- Main memory cleared daily
- Daily RDF files generated
Data Tagging

- Heterogeneous data usually problematic
- Two sensor types:
  - Temperature: thermocouples
  - Relative humidity: RH sensors
- Data are time stamped (has_date & has_time)

Ontology Representation

- Function 1: Constraints & Initialization
  - Grain/seed specific constraints
  - Indication of contents within mini-dome
  - Utilization of OWL
Function 2: Record Storage
- Predefined ontology schema
- Each record consists of 21 attributes
- `has_date` & `has_time` provide uniqueness
- Utilization of RDF/RDFS

Ontology Representation…
Query Processing

- Collaboration of SPARQL and Jena
- Support for three types of queries:
  - Exploratory
  - Monitoring
  - Range

Query Processing

- Needed files determined for query
- Files returned to main memory
- Files released upon completion of query
Semantic Benefits

- Providing meaning to meaningless data
- Exploit literal statements
- Query Richness

Future Work

- Addition of BRAHMS
  - Large RDF storage system
  - Supports fast semantic association discovery
  - Aid in data analysis
Conclusion

- Grain storage issues with Cargill & NPRL
- ES3N:
  - Data Acquisition
  - Data Storage
  - Data Management
  - Semantic Relief